AL/HR-TP-1995-0034

A PATH MODEL OF U.S. AIR FORCE PILOT TRAINING AND ITS ANTECEDENTS

Malcolm James Ree
Thomas R. Carretta

HUMAN RESOURCES DIRECTORATE
MANPOWER AND PERSONNEL RESEARCH DIVISION
7909 Lindbergh Drive
Brooks AFB, Texas 78235-5352

Mark S. Teachout

HUMAN RESOURCES DIRECTORATE
TECHNICAL TRAINING RESEARCH DIVISION
7909 Lindbergh Drive
Brooks AFB, Texas 78235-5352

December 1995

Interim Technical Paper for Period October 1994 - July 1995

ARMSTRONG LABORATORY

AIR FORCE MATERIEL COMMAND
BROOKS AIR FORCE BASE, TEXAS

NOTICE

Publication of this paper does not constitute approval or disapproval of the ideas or findings. It is published in the interest of scientific and technical information (STINFO) exchange.

When Government drawings, specifications, or other data are used for any purpose other than in connection with a definitely Government-related procurement, the United States Government incurs no responsibility or any obligation whatsoever. The fact that the Government may have formulated or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication, or otherwise in any manner construed, as licensing the holder, or any other person or corporation; or as conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

The Office of Public Affairs has reviewed this paper, and it is releasable to the National Technical Information Service, where it will be available to the general public, including foreign nationals.

This paper has been reviewed and is approved for publication.

Malcolm James Ree

MALCOLM JAMES REE
Scientific Advisor
Manpower & Personnel Research Div.

PATRICK C. KYLLONEN
Technical Director
Manpower & Personnel Research Div.

Gary D. Zank

GARY D. ZANK, Colonel, USAF
Chief, Manpower & Personnel Research Div.

Please notify this office, AL/HRPP, 7909 Lindbergh Drive, Brooks AFB TX 78235-5352, if your address changes, or if you no longer want to receive our techinical reports. You may write or call the STINFO office at DSN 240-3853 or commercial (210) 536-3853.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1995	Interim Technical Paper - October 1994-July 1995

4. TITLE AND SUBTITLE

Path Model of U.S. Air Force Pilot Training and Its Antecedents

5. FUNDING NUMBERS

PE - 62205F
PR - 7719
TA - 25
WU - 01

6. AUTHOR(S)

Malcolm J. Ree   Mark S. Teachout
Thomas R. Carretta

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Armstrong Laboratory
Human Resources Directorate
Manpower and Personnel Research Division (AL/HRM)
7909 Lindbergh Drive
Brooks AFB TX

8. PERFORMING ORGANIZATION REPORT NUMBER

AL/HR-TP-1995-0034

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING/MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

Also published in Journal of Applied Psychology, Vol. 80 (6), 1995.

12a. DISTRIBUTION/AVAILABILITY STATEMENT

Approved for public release; distribution is unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

A causal model of the role of general cognitive ability and prior job knowledge in subsequent job knowledge acquisition and work sample performance during training was developed. Participants were 3,428 Air Force officers in pilot training. The measures of ability and prior job knowledge came from the Air Force Officer Qualifying Test. The measures of job knowledge acquired during training were derivied from classroom grades. Work sample measures came from check flight ratings. The causal model showed that ability directly influenced the acquisition of job knowledge. General cognitive ability influenced work samples through job knowledge. Prior job knowledge had almost no influence on subsequent job knowledge, but directly influenced the early work sample. Early training job knowledge influenced subsequent job knowledge and work sample performance. Finally, early work sample performance strongly influenced subsequent work sample performance.

14. SUBJECT TERMS

Causal models   Pilot training
General cognitive ability
Job knowledge

15. NUMBER OF PAGES

24

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	UNLIMITED

CONTENTS

FIGURES

TABLES

Preface

This effort was performed under WU 77192501 in support of aircrew selection and classification research and development. The authors thank Winston Bennett, Anne Duke, James Earles, Lisa Follin, Mike Murray, and Paul Rioux for their assistance and contribution. The authors also thank the reviewers at *The Journal of Applied Psychology* for their many helpful comments on earlier drafts of this paper.

A PATH MODEL OF U.S. AIR FORCE PILOT TRAINING AND ITS ANTECEDENTS

Malcolm James Ree
Thomas R. Carretta
Mark S. Teachout

Summary

A causal model was developed that investigated the role of general cognitive ability (g) and prior job knowledge in subsequent job knowledge acquisition and work sample performance during sequential training. Participants were 3,428 U. S. Air Force officers enrolled in a 53 week pilot training program. The measures of ability and prior job knowledge came from the Air Force Officer Qualifying Test. The measures of job knowledge acquired during training were derived from pilot training classroom grades. Work sample job performance measures came from check flight ratings. The causal model showed that ability directly influenced the acquisition of job knowledge both prior to and during training. General cognitive ability influenced work sample performance indirectly through job knowledge, but did not show any direct influence. This may have occurred because all participants had a low level of job (flying) experience. Prior job knowledge had almost no influence on subsequent job knowledge, but directly influenced the early work sample. Early training job knowledge influenced subsequent job knowledge and work sample performance. Finally, early work sample performance strongly influenced subsequent work sample performance. Implications for the use of measures of g and prior job knowledge as personnel selection variables are discussed.

Introduction

Several studies have shown the effectiveness of general cognitive ability (g) as a predictor of numerous occupational criteria (e.g., McHenry, Hough, Toquam, Hanson, & Ashworth, 1990; McNemar, 1964; Ree & Earles, 1992). McNemar (1964), and Ree and Earles (1991) demonstrated that g accounted for almost all of the validity of multiple aptitude batteries for predicting training success, while McHenry et al. (1990) and Ree, Earles, and Teachout (1994) demonstrated that g was the best predictor of core technical job performance. Hunter (1986) provided a major summary article demonstrating that "...general cognitive ability has high validity predicting performance ratings and training success in all jobs." (p. 359)

Further, Olea and Ree (1994) extended these findings in a large scale study of about 5,400 U. S. Air Force (USAF) aviation trainees including traditional training criteria and hands-on work samples of flying performance. They found that job knowledge was incremental (.08) to the validity of g ($\underline{r}$ = .314) for pilots. Humphreys (1986) observed similar results during World War II.

Finally, Dye, Reck, and McDaniel (1993) demonstrated the validity of job knowledge tests for many jobs. They defined job knowledge as "...the cumulation of facts, principles, concepts, and other pieces of information that are considered important in the performance of one's job." (p. 153). In their meta-analysis of 502 validity coefficients based on 363,528 individuals, they found a mean validity of .47 for predicting training performance and .45 for predicting job performance.

In addition to the validity of g and job knowledge, their causal roles in job performance have been demonstrated. Hunter (1983) reported a path analysis of meta-analytically cumulated correlations relating g, job knowledge, and job performance. In the 14 studies with 3,264 participants he found that the major causal influence of cognitive ability was on the acquisition of job knowledge. Job knowledge, in turn, had a major impact on work sample performance and supervisory ratings. Work sample performance also directly influenced supervisory ratings. No direct effect of ability on supervisory job performance ratings was reported. In Hunter's model, work sample performance and job knowledge mediated (James & Brett, 1984) the relationship between ability and supervisory ratings.

Schmidt, Hunter, and Outerbridge (1986) extended Hunter (1983) to include experience on the job. They found that experience influenced measures of job knowledge and work sample measures of job performance. These latter two measures directly influenced supervisory ratings of job performance.

Borman, White, Pulakos, and Oppler (1991) confirmed Hunter's (1983) model in an additional sample of job incumbents, but went on to make it more parsimonious showing sequential effects from ability to job knowledge to task proficiency to supervisory ratings. They found that the paths from ability to task proficiency and from job knowledge to supervisory ratings were unnecessary. Borman et al. (1991) attributed this to the uniformity of job experience of the participants. More recently, Borman, White, and Dorsey (1995) confirmed the Borman et al. (1991) parsimonious model on peer and supervisory samples.

Borman, Hanson, Oppler, Pulakos, and White (1993) extended the model to supervisory job performance again showing that ability influenced job knowledge. They added job experience as a variable and significant paths between experience and proficiency, job knowledge, and supervisory ratings were found. Additionally, a path between ability and experience was found.

The current study investigated the causal role of g and prior job knowledge in a complex sequential training environment. Others have explored the role of g and job knowledge in job performance. We have extended this to training performance and make the distinction between prior job knowledge and job knowledge acquired during training. Specifically, we investigated a model of the causal role of g and prior job knowledge in acquiring subsequent job knowledge during training and the influence of g and job knowledge on work sample performance during training.

The hypotheses in the current study are based on previous findings (Borman et al., 1991, 1993, 1995; Hunter, 1983; Schmidt et al., 1986). The current study is like Borman et al. (1991, 1995) because the participants have similar levels of experience. In addition, Hunter (1983) and Schmidt, et al. (1986) suggested that the direct impact of g on performance in military samples would be small. They suggested this because of the highly structured nature of military job training. Based on all these previous findings, we hypothesized that g would have either weak or no direct influence on the work samples and that g would exert its influence indirectly through job knowledge. Further, as job knowledge is acquired as a consequence of g, prior job knowledge would be a direct function of g. Additional job knowledge should be a direct or indirect consequence of g. In the circumstance where

additional job knowledge relied heavily on preceding job knowledge, the effects of g were hypothesized to be indirect. Prior research suggested that there should be a causal link between job knowledge and work sample performance. Also, early work sample performance was expected to influence subsequent work sample performance because there is a sequential requirement to master one to advance to the other. Figure A1 shows the hypothesized model.

Method

Participants

The participants were 3,428 USAF officers that had attended and completed a 53 week pilot training course between the years 1981 and 1993. The attrition rate for this course varies little from the long term average of 22 %. Although the exact number starting the training program is not maintained in official records, we estimate that the size of the sample that began training aggregated across the years was about 4,400. The sample used for analyses was predominantly male (98.3%) and White (96.8%). Participants were between the ages of about 22 and 27 years and all had completed at least a baccalaureate degree from an accredited college or university. All participants had been selected for officer commissioning and for pilot training, in part, on the basis of their scores on the Air Force Officer Qualifying Test (AFOQT). The Air Force uses a selection board technique that rates applicants for admission into flying training on the basis of academic achievement with a preference toward science majors. Prior flying experience in some cases, medical fitness, and personal recommendations are also considered. These are not retained in archival files and were not available for inclusion in this study.

Measures

g and Prior Job Knowledge. The measures of g and prior job knowledge were extracted from the AFOQT, the paper-and-pencil test used to select students for officer precommissioning programs (i.e., Officer Training School, Reserve Officer Training Corps) and to qualify commissioned officers for pilot and navigator training programs. Like other military tests, the AFOQT is based on a detailed taxonomy of test and item specifications. This taxonomy defines the content (Berger, Gupta, Berger, & Skinner, 1990; Gupta, Berger, Berger, & Skinner, 1989) and psychometric properties (Skinner & Ree, 1987) of each test.

The AFOQT consists of 16 tests that measure g, flying job knowledge, and four lower-order cognitive factors: verbal, quantitative, spatial, and perceptual speed (Carretta & Ree, 1996). For purposes of this study, verbal and quantitative tests, the most universally accepted measures of general cognitive ability, were used to estimate g. Prior job knowledge (JK_p) was assessed through measures of instrument comprehension and aviation information. Descriptions of the tests grouped by content are provided below.

Verbal tests. Verbal Analogies (VA) measures the ability to reason and recognize relationships between words. Reading Comprehension (RC) assesses the ability to read and understand paragraphs.

Word Knowledge (WK) provides a measure of ability to understand words through the use of synonyms.

Quantitative tests. Arithmetic Reasoning (AR) measures the ability to understand arithmetic relationships expressed as word problems. Data Interpretation (DI) assesses the ability to extract information from data presented in tables and charts. Math Knowledge (MK) requires the ability to use mathematical terms, formulas, and relationships to solve problems. Scale Reading (SR) measures the ability to read and extract information from scales and dials.

Prior job knowledge tests. These are the only tests in the AFOQT that measure specific job knowledge (Dye et al., 1993; Olea & Ree, 1994). Instrument Comprehension (IC) assesses the ability to determine the attitude (position and orientation in three dimensional space) of an aircraft in flight based on information presented in illustrations of flight instruments. Aviation Information (AI) measures knowledge of general aviation principles, concepts, and terminology.

Pilot Academic and Flying Grades

Pilot academic grades. Academic indicators represented student pilots' performance on written tests of flying theory, procedures, and aircraft-unique systems (i.e., hydraulics, instruments, electronics, etc.) learned during training. On each academic test, students received a percent correct score. There were 11 end-of-course tests (A1 through A11). They were divided into three groups that represented early (A1 to A4), middle (A5 to A8), and late training (A9 to A11). Early and middle classroom training were relevant to flying the subsonic primary aircraft (T-37). Early classroom training included courses in T-37 systems, aerospace physiology/human factors, flying fundamentals, and T-37 aerodynamics. Middle classroom training provided courses relevant to flight in general and to flying the primary aircraft and included T-37 instruments I and II, T-37 navigation, and T-37 mission planning. Late classroom training was relevant to the supersonic advanced aircraft (T-38) including T-38 systems operations, applied aerodynamics, and flight planning.

Flying work samples. There are two general categories of flights during training that accumulate about 190 flying hours. On ordinary daily flights the student pilot learns and practices under the watchful eye of an instructor pilot. After the prescribed daily flights, work sample tests called check flights are rated by check flight pilots. To eliminate potential bias due to familiarity, check flight pilots do not rate students with whom they have flown on daily flights.

During training, student pilots completed three check flights in the primary aircraft (CF1 to CF3) and three in the advanced aircraft (CF4 to CF6). In the primary aircraft, students must (1) demonstrate the ability to fly to a geographical area, perform maneuvers and return to perform successful landings, (2) conduct airborne activities within precise altitude and geographical limits, and (3) use instruments with an emphasis on landing approaches.

The advanced training aircraft is much faster than the primary training aircraft and all activities must be accomplished more rapidly. This makes even familiar maneuvers more difficult in the advanced aircraft. The check flights for round trips to geographical areas and instruments are similar to

the check flights in the primary aircraft. The difficult formation check flight is added where the wings of multiple aircraft are as close as three feet at speeds of 400 knots. See Duke and Ree (in press) for a complete description.

Each check flight grade score was a weighted average of ratings of several flying procedures and maneuvers. These procedures, maneuvers, and weights are prescribed by the Air Force in training regulations. The student pilot receives points for each procedure accomplished. Example procedures are: retract landing gear at specified speed, make proper radio calls during flight, or perform a loop within specified parameters (e.g., engine power settings and maneuver entry altitude). As with academic grades, check flight grades were percentage scores.

The sequential training environment was structured as follows. Theory and general background were taught first in the classroom followed by application in the aircraft. Classroom training for the primary aircraft began prior to check flight work samples. The final check flight work sample in the primary aircraft was completed after the last classroom instruction in middle training (A5 to A8). After the final check flight in the primary aircraft, classroom instruction relevant to the advanced aircraft began and was followed, shortly thereafter, by the check flight work samples. The last check flight work sample in the advanced aircraft occurred after all the classroom training had been completed.

Procedures

Measures of g and job knowledge acquired prior to training were included as were sequentially-ordered blocks of classroom training and hands-on work sample performance measures. Because the participants had been selected, at least in part, on the basis of the scores on the test battery that yielded the estimates of g and prior job knowledge, they constituted a censored, range restricted sample. Such samples provide relatively poor statistical estimates of the correlations among variables (Thorndike, 1949). We used the multivariate method (Lawley, 1943; Ree, Carretta, Earles, & Albert, 1994) to correct the correlation matrix for the consequences of prior selection. The corrected matrix was used in the structural equation analyses.

The Bentler-Weeks (Bentler & Weeks, 1980) structural model was estimated using maximum likelihood procedures as implemented in version 4.02 of the EQS program. This program corrects for unreliability via estimation in the same fashion as LISREL and most other structural modeling programs. The estimated reliabilities can either be provided as starting values or they can be estimated directly from the data. In these analyses, they were estimated directly from the data.

We followed the rules offered by Marsh (1994) as a general approach to evaluating goodness-of-fit. These are: (1) determine that the iterative procedure converged to a proper solution, (2) establish that parameter estimates were reasonable in accordance with the a priori model and common sense, and (3) evaluate the various fit indexes in relation to rules of thumb and values from competing models.

We considered several goodness-of-fit statistics and chose the Comparative Fit Index (CFI; Bentler, 1989) that has been shown to have little sample size dependence and a small sampling variance. The model χ^2 and root mean square error of approximation (RMSEA) also are presented. Finally, the residuals were inspected to determine their magnitude and to determine if any variable was being poorly predicted by the model.

We first fit the measurement model and then the path (causal) models. Because the scales of the variables are not well known or intrinsically meaningful, we reported the path coefficients as standardized regression coefficients (Cohen & Cohen, 1983). These coefficients should be interpreted as indicating that a standard deviation change in an independent variable leads to a change in the dependent variable equal to the magnitude of the path coefficient.

The causal model (Figure A2) based on Hunter (1983) and Schmidt et al. (1986) with all the hypothesized links was estimated. It was compared with a revised model (Figure A3) suggested by the findings of Borman et al. (1991). We computed the squared multiple correlations (R^2) for dependent variables for dependent variables.

Results

Tables 1 and 2 show the observed and unrestricted means, standard deviations, reliabilities, and correlations for all variables.

The measurement model is shown in Table 3. Loadings of the tests and other measurement variables can be found in this table. The aptitude tests loaded on lower-order verbal (V) and quantitative (Q) factors with g in hierarchical position. The specific job knowledge tests lead to a prior job knowledge (JK_p) factor. Job knowledge acquired during training was represented by factors derived from early (JK_{T1}), middle (JK_{T2}), and late (JK_{T3}) training as measured by job knowledge tests.

Two work sample performance factors (WS_1 and WS_2) were derived from the six check flights, three in each aircraft type. Latent factor correlations among g, prior job knowledge, job knowledge acquired during training, and work samples are shown in Table 4.

The measurement model fit the data well. The CFI was .969 which is very close to the maximum possible value of 1. The model χ^2 was 1,075.267 (df = 271) and was appreciably smaller than the independence χ^2 of 26,277.069 (df = 325). The RMSEA was .034 which Browne and Cudeck (1993) place in the "good" category. Additionally, the average absolute standardized residual was only .019, indicating good fit.

The hypothesized path model based on Hunter (1983) and Schmidt et al. (1986) was computed. Its estimated path coefficients are shown in Figure A2. The fit was good. The CFI was .968, the RMSEA was .034, and the average absolute standardized residual was .020. However, there were several non-significant paths ($g \rightarrow WS_1$, $g \rightarrow WS_2$, $g \rightarrow JK_{T2}$, $JK_p \rightarrow JK_{T1}$, and $JK_{T3} \rightarrow WS_2$). Figure A3 presents the revised model and its path coefficients. Based on the findings of Hunter (1983) and

Table 1. Means and Standard Deviations for AFOQT Tests, Academic Grades and Check Flight Grades

Score[a]	Restricted (Observed) Mean	SD	Unrestricted Mean	SD	Reliability[b]
VA	15.29	3.36	13.36	4.23	.80
AR	13.54	4.11	11.00	4.40	.81
RC	17.44	4.73	15.83	5.93	.88
DI	13.52	3.91	11.15	3.93	.71
WK	13.96	5.17	13.28	5.83	.88
MK	18.01	4.63	14.48	6.04	.88
SR	24.23	5.54	20.07	6.73	.84
IC	13.66	4.23	8.82	4.76	.84
AI	11.74	4.26	8.65	4.08	.77
A1	97.46	3.11	96.60	3.19	.23
A2	97.18	3.32	96.52	3.38	.20
A3	97.04	3.37	96.40	3.41	.18
A4	98.06	3.28	97.19	3.36	.22
A5	95.99	4.80	95.01	4.89	.24
A6	95.18	5.32	94.28	5.37	.20
A7	94.77	5.36	93.64	5.43	.20
A8	95.88	4.55	95.11	4.60	.19
A9	97.37	3.32	96.76	3.38	.12
A10	97.30	3.63	96.64	3.70	.21
A11	96.83	3.69	96.04	3.76	.20
CF1	86.59	7.56	85.12	7.60	.23
CF2	90.65	5.73	89.59	5.77	.24
CF3	93.59	4.87	92.61	4.93	.26
CF4	91.21	5.70	90.07	5.74	.18
CF5	92.65	4.67	91.89	4.69	.15
CF6	93.82	4.73	92.89	4.78	.17

Notes. [a]The first 9 variables are tests from the AFOQT: VA = Verbal Analogies, AR = Arithmetic Reasoning, RC = Reading Comprehension, DI = Data Interpretation, WK = Word Knowledge, MK = Math Knowledge, SR = Scale Reading, IC = Instrument Comprehension, AI = Aviation Information. A1 through A11 are the 11 job knowledge tests and CF1 through CF6 are the 6 check flight ratings.
[b]Reliabilities for the AFOQT tests were taken from Sinner and Ree (1987), the sample to which the data were corrected for range restriction. Reliabilities for the job knowledge scores and check flight ratings were the values estimated by EQS in the measurement model.

Table 2. Correlation[a] of AFOQT Tests, Academic Grades and Check Flight Grades

Score[b]	VA	AR	RC	DI	WK	MK	SR	IC	AI	A1	A2	A3	A4	A5	A6	A7	A8	A9	A10	A11	CF1	CF2	CF3	CF4	CF5	CF6
VA	100	48	58	47	60	39	33	27	17	16	17	14	14	12	08	12	11	14	13	12	01	05	06	01	03	05
AR	58	100	44	63	42	57	55	30	12	17	14	17	16	16	14	18	14	14	19	15	07	08	09	06	04	08
RC	73	58	100	43	69	38	32	23	21	17	16	14	16	11	09	12	12	15	15	12	00	03	06	02	01	06
DI	53	67	55	100	44	43	51	35	21	16	12	14	14	12	12	19	14	12	14	16	08	10	10	08	06	10
WK	68	46	77	46	100	33	29	26	27	13	15	12	11	06	07	10	10	11	11	08	00	02	03	-01	00	02
MK	55	71	51	60	40	100	42	18	00	18	15	16	17	16	12	15	12	16	22	18	05	07	12	07	05	12
SR	48	66	45	62	37	60	100	36	18	14	10	12	11	14	11	13	14	11	15	15	10	11	13	10	06	13
IC	34	41	33	43	28	39	49	100	45	12	09	09	11	05	05	07	06	06	05	06	10	09	08	10	09	08
AI	30	31	34	34	32	25	33	56	100	13	10	09	11	03	07	10	08	07	03	05	08	05	04	03	02	01
A1	24	26	24	25	18	28	23	19	20	100	19	19	23	20	18	14	13	19	19	14	07	09	12	06	04	08
A2	25	22	23	19	20	23	18	15	16	23	100	15	21	19	17	18	17	16	15	17	05	07	07	06	05	08
A3	19	23	19	19	16	22	18	14	14	22	18	100	17	21	19	16	15	20	14	13	07	06	14	05	05	06
A4	23	25	24	23	17	27	21	19	19	27	24	21	100	12	20	19	13	19	16	12	09	08	10	10	04	07
A5	20	23	18	19	11	16	22	12	10	24	22	23	16	100	21	21	24	23	18	17	14	11	16	07	09	12
A6	14	19	14	17	10	13	17	10	11	20	20	21	22	23	100	18	19	23	18	18	13	14	12	10	08	12
A7	17	23	17	23	12	15	19	12	15	17	20	19	22	24	20	100	20	23	14	20	11	10	10	09	08	09
A8	18	20	18	19	14	12	20	11	12	16	20	17	16	27	21	22	100	22	17	18	10	12	10	08	06	07
A9	22	22	22	20	17	17	19	12	13	23	19	22	23	26	25	26	24	100	21	20	10	11	10	10	04	09
A10	20	27	21	21	15	20	22	12	10	23	18	17	20	21	20	17	19	24	100	20	12	13	11	13	07	14
A11	19	23	18	23	12	18	23	14	12	18	20	16	16	20	20	22	20	23	24	100	12	11	13	09	05	11
CF1	03	10	02	11	00	08	14	13	10	09	06	08	11	16	14	12	11	11	14	14	100	25	25	21	21	17
CF2	09	13	07	14	04	11	16	13	09	11	09	08	10	13	15	12	13	13	15	13	26	100	28	19	19	17
CF3	12	15	11	16	06	10	19	13	09	15	10	16	13	18	14	12	13	13	14	15	26	29	100	19	19	19
CF4	05	11	05	13	00	12	15	14	08	09	08	07	12	09	12	11	09	12	15	11	22	20	20	100	19	16
CF5	06	07	03	09	01	07	10	12	05	06	06	06	06	11	09	10	07	06	09	07	22	20	20	19	100	15
CF6	11	15	11	16	06	09	18	13	06	11	11	09	10	15	14	12	09	12	16	14	19	18	21	18	16	100

Notes. [a]Correlations above the diagonal are observed data. Those below the diagonal were corrected for range restriction (Lawley, 1943).
[b]The first 9 variables are tests from the AFOQT: VA = Verbal Analogies, AR = Arithmetic Reasoning, RC = Reading Comprehension, DI = Data Interpretation, WK = Word Knowledge, MK = Math Knowledge, SR = Scale Reading, IC = Instrument Comprehension, AI = Aviation Information. A1 through A11 are the 11 job knowledge tests and CF1 through CF6 are the 6 check flight ratings.

Table 3. Factor Loadings of the Measurement Model

	Factor[b]								
Score[a]	V	Q	g	JK_P	JK_{T1}	JK_{T2}	JK_{T3}	WS_1	WS_2
VA	.46		.68						
AR		.51	.83						
RC	.60		.67						
DI		.03	.61						
WK	.68		.54						
MK		.13	.77						
SR		.05	.76						
IC				.86					
AI				.65					
A1					.51				
A2					.46				
A3					.44				
A4					.50				
A5						.52			
A6						.48			
A7						.47			
A8						.46			
A9							.52		
A10							.48		
A11							.47		
CF1								.51	
CF2								.52	
CF3								.55	
CF4									.44
CF5									.41
CF6									.43

Notes. [a]The first 9 variables are tests from the AFOQT: VA = Verbal Analogies, AR = Arithmetic Reasoning, RC = Reading Comprehension, DI = Data Interpretation, WK = Word Knowledge, MK = Math Knowledge, SR = Scale Reading, IC = Instrument Comprehension, AI = Aviation Information.
A1 through A11 are the 11 job knowledge tests and CF1 through CF6 are the 6 check flight ratings.
[b]The factors are V = Verbal, Q = Quantitative, g = general cognitive ability, JK_P = prior job knowledge, JK_{T1} = job knowledge acquired during training (measure 1), JK_{T2} = job knowledge acquired during training (measure 2), JK_{T3} = job knowledge acquired during training (measure 3), WS_1 = flying training work sample (measure 1), and WS_2 = flying training work sample (measure 2).

Table 4. Correlations Between Factors[a] in the Causal Model

Factor	g	JK_P	JK_{T1}	JK_{T2}	JK_{T3}	WS_1	WS_2
g	1.00						
JK_P	.62	1.00					
JK_{T1}	.63	.45	1.00				
JK_{T2}	.55	.30	.87	1.00			
JK_{T3}	.61	.33	.86	.95	1.00		
WS_1	.33	.29	.44	.56	.54	1.00	
WS_2	.37	.35	.44	.54	.56	.92	1.00

Note. [a]The factors are V = Verbal, Q = Quantitative, g = general cognitive ability, JK_P = prior job knowledge, JK_{T1} = job knowledge acquired during training (measure 1), JK_{T2} = job knowledge acquired during training (measure 2), JK_{T3} = job knowledge acquired during training (measure 3), WS_1 = flying training work sample (measure 1), and WS_2 = flying training work sample (measure 2).

Table 5. Variance Accounted for in the Dependent Variables

	R^2	
Variable[a]	Initial Model	Revised Model
JK_p	.38	.38
JK_{T1}	.40	.40
JK_{T2}	.74	.74
JK_{T3}	.94	.94
WS_1	.32	.31
WS_2	.86	.87

Note. [a]The factors are JK_P = prior job knowledge, JK_{T1} = job knowledge acquired during training (measure 1), JK_{T2} = job knowledge acquired during training (measure 2), JK_{T3} = job knowledge acquired during training (measure 3), WS_1 = flying training work sample (measure 1), and WS_2 = flying training work sample (measure 2).

Schmidt et al. (1986) the direct paths from g to the work samples were removed. There was no theoretical basis for removing the other non-significant paths and they were retained. All of the fit indexes stayed the same as in the initial model, indicating no loss in fit due to removal of paths. The proportion of variance accounted for (R^2) for each dependent variable for both causal models is presented in Table 5.

Discussion

We conducted a longitudinal study to investigate the causal roles of g and prior job knowledge in a complex, sequential training environment. While the initial model fit the data quite well, some paths were non-significant. The revised model required fewer parameters and showed no decrease in fit. Additionally, it was closer to a model proposed by Borman et al. (1991, 1995). The reason the revised model was more appropriate for these data is probably due to the thoroughness of detailed military training programs and strict adherence to standard operating procedures (Hunter, 1983; Schmidt et al., 1986). In the military, performance requirements are laid out in great detail in technical manuals and training programs. Hence, in military samples, ability is less likely to have a direct impact on performance. It is more likely that ability will have an indirect impact on performance through job knowledge as found in the current study.

In the revised model, g was found to influence work sample performance only indirectly through prior job knowledge and job knowledge acquired during training. This was consistent with and extends past research that has shown the relationship of g to job performance (Borman et al., 1991, 1993, 1995; Hunter, 1983; Schmidt et al., 1986).

In the revised model, g led directly to the acquisition of prior job knowledge (JK_P) and job knowledge acquired during early (JK_{T1}) and late training (JK_{T3}). Surprisingly, the direct path from g to middle job knowledge (JK_{T2}) was near zero. The causal influence of g on JK_{T2} was mediated by JK_P and JK_{T1}.

A direct path from g to work sample performance was not necessary. The influence of g on work samples was entirely mediated by job knowledge. This was expected because g leads to job knowledge and job knowledge leads to job performance, particularly in military contexts where standard operating procedures are detailed in technical manuals and training programs.

Job knowledge acquired in early training influenced subsequent job knowledge acquisition and work sample performance. Early training job knowledge generally dealt with the basics, while later job knowledge added information about specific aircraft systems (hydraulics, instruments, unique landing requirements, etc.). The course content progression was from general to specific. The paths from JK_{T1} to JK_{T2} and from JK_{T2} to JK_{T3} were about equal and quite strong. Some of the magnitude of these paths may have come from the fact that early courses are foundations for subsequent courses. In JK_{T1}, the course in aerodynamics is a prerequisite for the applied aerodynamics course in JK_{T3}. Also in JK_{T1}, the systems course is prerequisite to the instruments I and II courses in JK_{T2}. Similarly, the navigation course and mission planning course in JK_{T2} are prerequisites for the flight planning course in JK_{T3}.

Finally, flying fundamentals in JK_{T1} is a building block for several of the courses in both JK_{T2} and JK_{T3}.

The combined direct and indirect paths from job knowledge to early work sample performance (WS_1) were about equal to those found in previous research when experience was held constant. The direct path of the late training job knowledge (JK_{T3}) to WS_2 was small. The indirect paths through WS_1 were much larger.

The single direct path from WS_1 to WS_2 was very strong indicating a powerful causal role for skills and experience acquired in early samples of job performance. Part of the strength of this path may have been because the activities learned flying the primary aircraft were required in flying the advanced aircraft. Both required preflight preparation, take-offs, aeronautic maneuvers, landings, and other elements.

In this study there was no final dependent measure equivalent to the supervisory ratings reported by Hunter (1983), Schmidt et al. (1986), and Borman et al. (1991, 1993, 1995). However, WS_1 is comparable to the variable called "work sample" in Hunter (1983) and in Schmidt et al. (1986), "task proficiency" in Borman et al. (1991), "supervisory proficiency" in Borman et al. (1993), and "technical proficiency" in Borman et al. (1995).

We have demonstrated the causal role of general cognitive ability and prior job knowledge in a complex training environment. The causal role of g is through the acquisition of job knowledge that in turn has an impact on work sample performance. The causal impact of prior job knowledge was very weak. Its influence on early work sample performance was greater than its influence on job knowledge acquired during training. Job knowledge acquired during training had a strong impact on job knowledge acquired later in training, but only moderate to weak influence on work sample performance. Not surprisingly, learning to fly the primary aircraft was highly influential in learning to fly the advanced aircraft.

Selecting applicants with high scores on g and high scores on prior job knowledge should lead to better performance in training and thereby better performance on the job as shown in previous research (Borman et al., 1991, 1993; Hunter, 1983; Olea & Ree, 1994; Schmidt et al. 1986, 1988). However, the causal influence of prior job knowledge was very small compared to that of g. This suggests that given competition for resources, testing of general cognitive ability might yield greater gains and should take precedence over testing for prior job knowledge.

References

Bentler, P. M. (1989). *EQS structural equations program manual.* Los Angeles: CA, BMDP Statistical Software.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*, 238-246.

Bentler, P. M., & Weeks, D. G. (1980). Linear structural equations with latent variables. *Psychometrika, 45*, 289-308.

Berger, F. R., Gupta, W. B., Berger, R. M., & Skinner, J. (1990). *Air force officer qualifying test (AFOQT) form P test manual* (AFHRL-TR-89-56). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Borman, W. C., White, L. A., & Dorsey, D. W. (1995). Effects of ratee task performance and interpersonal factors on supervisor and peer performance ratings. *Journal of Applied Psychology, 80*, 168-177.

Borman, W. C., White, L. A., Pulakos, E. D., & Oppler, S. H. (1991). Models of supervisory job performance ratings. *Journal of Applied Psychology, 76*, 863-872.

Borman, W. C., Hanson, M. A., Oppler, S. H., Pulakos, E. D., & White, L. A. (1993). Role of early supervisory experience in supervisor performance. *Journal of Applied Psychology, 78*, 443-449.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen and J. S. Lang (Eds.) *Testing Structural Equation Models*. Newbury Park, CA: Sage.

Carretta, T. R., & Ree, M. J. (1995). Air Force Officer Qualifying Test validity for predicting pilot training performance. *Journal of Business and Psychology, 9*, 379-388.

Carretta, T. R., & Ree, M. J. (1996). Factor structure of the Air Force Officer Qualifying test: Analysis and comparison. *Military Psychology, 8*, 29-42.

Cohen, J. & Cohen, P. (1983). *Applied multiple regression/correlation analyses for the behavioral sciences.* (2nd ed.). Hillsdale, NJ: Erlbaum.

Duke, A. P., & Ree, M. J. (in press). *Better candidates fly fewer training hours: Another time testing pays off.* (AL/HR-TP-1995-XX) Brooks AFB TX: Manpower and Personnel Research Division, Armstrong Laboratory.

Dye, D. A., Reck, M., & McDaniel, M. (1993). The validity of job knowledge measures. *International Journal of Selection and Measurement, 1*, 153-162.

Gupta, W. B., Berger, F. R., Berger, R. M., & Skinner, J. (1989). *Air Force Officer Qualifying Test (AFOQT): Development of an item bank* (AFHRL-TP-89-33). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Humphreys, L. G. (1986). Commentary. *Journal of Vocational Behavior, 29*, 241-247.

Hunter, J. E. (1983). A causal analysis of cognitive ability, job knowledge, job performance, and supervisor ratings. In F. Landy, S. Zedeck, & J. Cleveland, *Performance Measurement and Theory*, Hillsdale, New Jersey: Erlbaum. 257-266.

Hunter, J. E. (1986). Cognitive ability, cognitive aptitudes, job knowledge, and job performance. *Journal of Vocational Behavior, 29*, 340-362.

James, L. R., & Brett, J. M. (1984). Mediators, moderators, and tests for mediation. *Journal of Applied Psychology, 69*, 307-321.

Lawley, D. N. (1943). A note on Karl Pearson's selection formulas. *Proceedings of the Royal Society of Edinburgh, Section A, 62*, Part I, 20-30.

Marsh, H. W. (1994). Confirmatory factor analysis models of factorial invariance: A multifaceted approach. *Structural Equation Modeling, 1*, 5-34.

McHenry, J. J., Hough, L. M, Toquam, J. L., Hanson, M., & Ashworth, S. (1990). Project A validity results: The relationship between predictor and criterion domains. *Personnel Psychology, 43*, 335-354.

McNemar, Q. (1964). Lost: our intelligence? Why? *American Psychologist, 19*, 871-882.

Olea, M., & Ree, M. J. (1994). Predicting pilot and navigator criteria: Not much more than g. *Journal of Applied Psychology, 79*, 845-851.

Ree, M. J., Carretta, T. R., Earles, J. A. , & Albert, W. (1994). Sign changes when correcting for range restriction: A note on Pearson's and Lawley's selection formulae. *Journal of Applied Psychology, 79*, 298-301.

Ree, M. J., & Earles, J. A. (1991). Predicting training success: Not much more than g. *Personnel Psychology, 44*, 321-332.

Ree, M. J., & Earles, J. A. (1992). Intelligence is the best predictor of job performance. *Current Directions in Psychological Science, 1*, 86-89.

Ree, M. J., & Earles, J. A. (1993). g is to psychology what carbon is to chemistry. A reply to Sternberg and Wagner, and to McClelland and Calfee. *Current Directions in Psychological Science, 2*, 11-12.

Ree, M. J., Earles, J. A., & Teachout, M. S. (1994). Predicting job performance; Not much more than g. *Journal of Applied Psychology, 79*, 518-524.

Schmidt, F. L., Hunter, J. E., & Outerbridge, A. N. (1986). Impact of job experience and ability on job knowledge, work sample performance, and supervisory ratings of job performance. *Journal of Applied Psychology, 71*, 432-439.

Schmidt, F. L., Hunter, J. E., Outerbridge, A. N., & Goff, S. (1988). Joint relation of experience and ability with job performance: Test of three hypotheses. *Journal of Applied Psychology, 73,* 46-57.

Skinner, J., & Ree, M. J. (1987). *Air Force Officer Qualifying Test (AFOQT): Item and factor analysis of Form O* (AFHRL-TR-86-68). Brooks AFB, TX: Manpower and Personnel Division, Air Force Human Resources Laboratory.

Thorndike, R. L. (1949). *Personnel Selection*, NY: Wiley.

APPENDIX

Three Models for Sequential Training

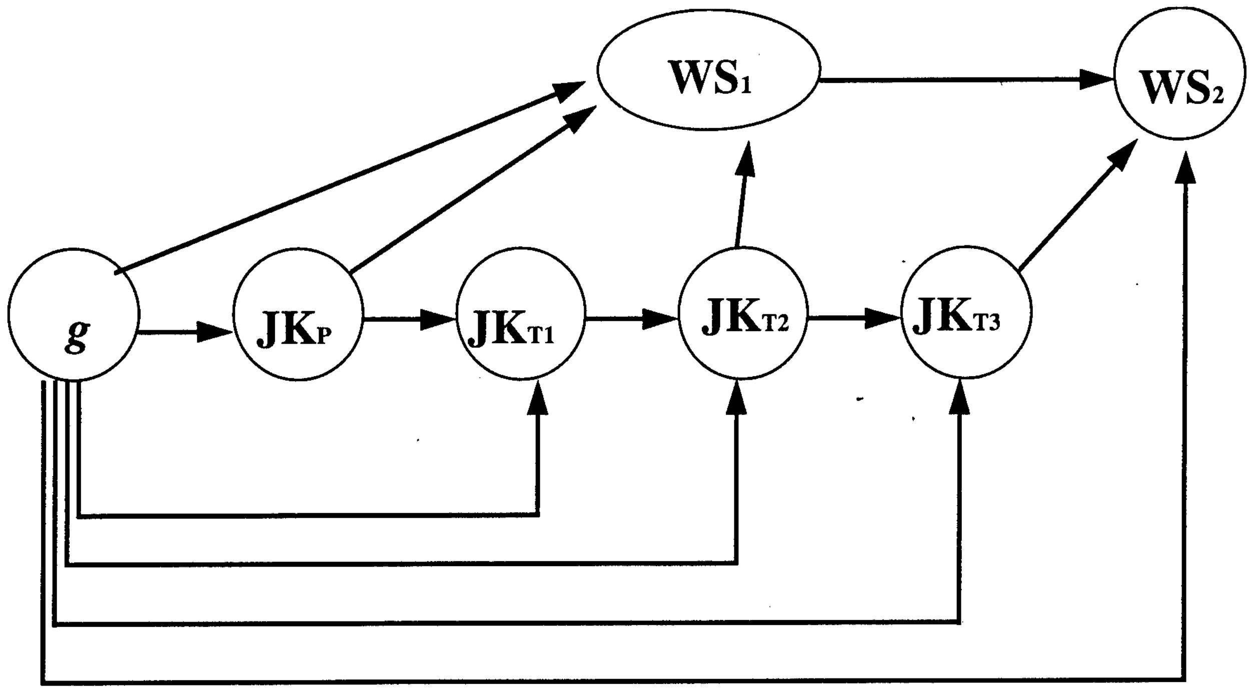
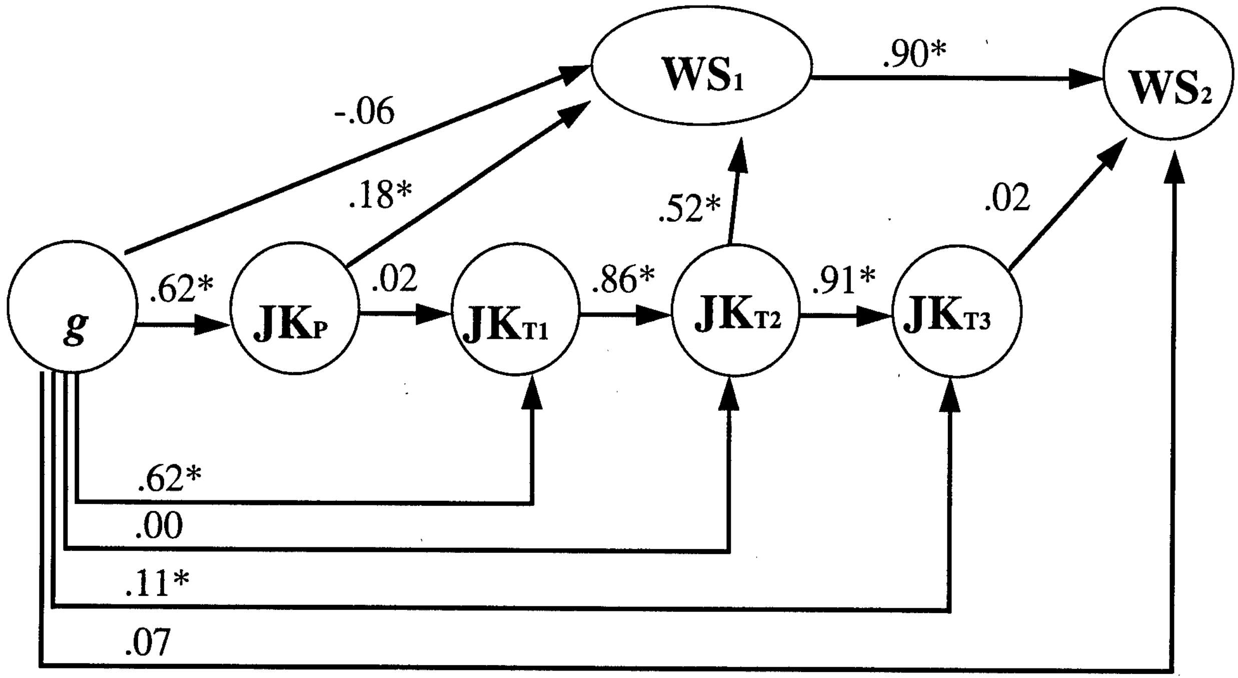
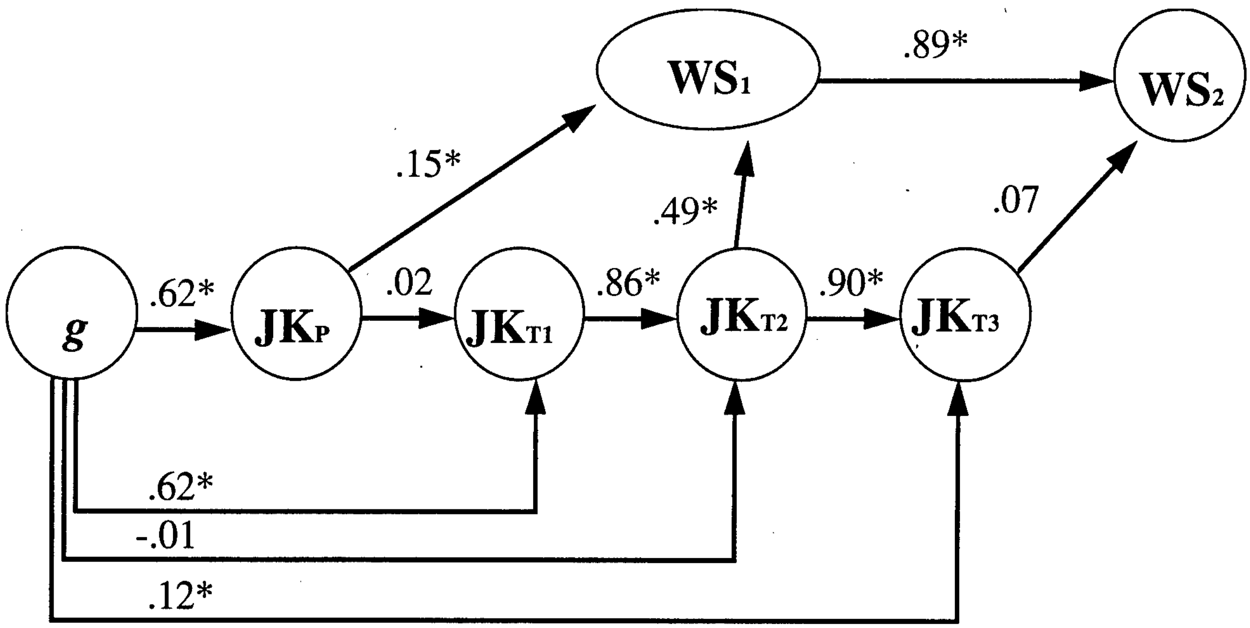


Figure A1. Hypothesized Causal Model for Sequential Training



*p < .01

Figure A2. Initial Causal Model and Path Coefficients for Sequential Training



*p < .01

Figure A3. Revised Causal Model and Path Coefficients for Sequential Training